PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	Clinical specialty training in UK undergraduate medical schools: a retrospective observational study
AUTHORS	Vaidya, Hrisheekesh; Emery, Alexander; Alexander, Emma; McDonnell, Angus; Burford, Charlotte; Bulsara, Max

VERSION 1 - REVIEW

REVIEWER	Charles Weissman Hadassah Hebrew University Medical Center, Israel
REVIEW RETURNED	05-Aug-2018

GENERAL COMMENTS	This manuscript examines the relationship of time spent in the various clinical rotations during medical school and the residency programs the students entered. 1. The premise of this manuscript is somewhat simple. Namely, that the amount of clinical exposure students get in medical school will direct them to select the specialty for residency, i.e. the more tthe exposure the more likely the students will select the specialty as a career. Is this true? It is much more complicated, since it is also the quality of the experience and the presence of influential mentors and the enthusiasm generated toward the students and the specialty during their rotation which are influential in specialty selection (there are many papers written on the subject of mentorsip, positive/negative experiences and specialty selection). Furthermore, personal and socio-economic factors play important roles in specialty selection and were not examined in this study. 2. The role of the medical school is to provide all graduates with a good foundation in clinical medicine no matter which specialty the student will enter. Therefore, exposure during the clinical rotations are geared to providing the student with a "bread and butter" experience. It is not the main mission of the medical school to direct students to specialties with workforce deficits. The medical school can help direct students by providing required exposures to these specialties but not at the expense of their core education. This issue needs to be addressed in the manuscript: Educational goals versus operational (healthcare system) goals. 3. If there is such a need for General Practitioners in the UK it is the obligation of the healthcare system to examine the problem and develop solutions as to why there is a problem. The medical schools can only be part of a comprehenisve solution since medical students are very savy as to the working conditions and remuneration of the various specialties. The authors did not examine the students' percentions, goals and decision making to

	examine the disconnect between the time spent in a specialty in medical school and their choices of specialties. This is a major weakness of this manuscript. What does the student think? 4. How have other countries dealt with these issues, specialties with workfore shortages versus medical school curriculae? Research has been performed on recruiting more students to general/family practice and to rural areas by gearing medical school curriculae and experiences to shed a positive light on this area of medical practice. These studies should be reviewed in the manuscript. 5. The authors include the surgical experience during medical school in their analysis, was it all general surgery or also exposure to surgical subspecialties? How was a student choosing urology or plastic surgery considered?

REVIEWER	Katherine Woolf UCL, UK Educational advisor MRCP(UK), National Institute for Health Research Fellow.
REVIEW RETURNED	14-Aug-2018

GENERAL COMMENTS	This paper has the potential to make a significant contribution to the literature on medical school training and specialty outcomes. I was particularly impressed at the efforts made to collect data. The paper needs much more clarity on the methods and results. Some of the statistical analysis may need redoing since I'm not certain that the non-statistically significant findings are not due to lack of power and/or outliers. Furthermore, I could not understand one of the analyses. Major comments: 1. The interpretation of the non-significant correlation between median number of weeks spent in each specialty and number of training posts is problematic. Because the analysis is done at the level of the specialty, there are only 12 data points (as can be seen in Fig 4). The correlation coefficient is 0.43, but with an alpha of 0.05 and a sample size of 12, the power is only 0.27 (calculations done with this online calculator https://www.anzmtg.org/stats/PowerCalculator/PowerCorrelation). I personally don't think statistical significance is particularly relevant here. To me the interesting thing is that GP is such a clear outlier, with many more posts than median weeks across medical schools. If we are going to think about a statistical relationship, the authors need to consider that such a large outlier is likely to significant skew the correlation, particularly with a small sample size. My guess from looking at the numbers in Fig 4 is that if GP were removed, the correlation may be much higher. One suggestion would be for the authors to keep Figure 4, but plot two fit lines, one including GP and one excluding it. This would allow the reader to have a visual representation of the specialties with relatively more weeks per post than average, and fewer weeks per post than average. I would not necessarily calculate the correlation; however if the authors really want to do it, I think they should remove GP as a very significant outlier and/or calculate a non-parametric correlation coefficient. 2. I cannot understand how the analysis looking at the relationship between the number of weeks per specialty and the percentage of

	entrants per specialty for each medical school was done. I can't understand what the dependent variable is. On p12 in the results section, I don't understand the sentence on line 252-253: what does "based on a univariate analysis of variance of our dataset" mean? How was specialty "a confounder"? More information is needed on p10 in the methods. This needs to be rewritten with more details before I can assess whether the analysis was appropriate, and how to interpret the findings. 3. In general the authors need to ensure consistency in the terms use to describe variables throughout the manuscript. On p10 it's stated that one of the variables is "the number of applicants from that medical school applying to each of the CT1/ST1 specialties". But then on p12 it is stated that the variable is in fact the "percentage of graduates from a medical school picking a specialty" (which is more appropriate considering medical schools produce different numbers of specialties). However on p8 it appears that these data are actually entrants to a specialty ("doctors directly entering a specialty training programme after foundation training") rather than applicants. Another example: on p7 an aim is to "examine whether the percentage of time spent in the different specialties correlated with the number of posts available at CT1/ST1". Then on p10 it is stated that "linear correlation was used to compare the median weeks spent in a specialty at medical school with […] the number of CT1/ST1 posts in 2016". The number of weeks spent is not the same as the percentage of time spent. On p12 the authors write "we found no correlation between the number of applicants to a specialty training programme…" Do they mean applications? Or did they consider people making multiple applications to different specialties? 4. The authors need to be clearer about the level at which each analysis is done. For example, the analysis above (correlation between median weeks per specialty and number of posts) is done at the level of the specialty, i.e. not taking into account variability between medical schools. The analysis of weeks per specialty and percentage of graduates entering the specialty is presumably done at the level of the medical school? Minor comments: 5. Descriptive statistics for each variable are needed. Figure 2 should be number in the text. I would find it helpful to have a table with the number and proportion of graduates per specialty per medical school. 6. Those not familiar with UK medical training would probably find it helpful to have a description, and an explanation of what ST1/CT1 means. 7. The authors asked about weeks during "clinical education", however in some schools the division between pre-clinical and clinical is not clear, in particular some less traditional schools have quite a lot of early patient contact. This needs to be mentioned.

REVIEWER	Hugh Alberti Newcastle University, UK None declared. I was the author of one of the papers that the authors quote and discuss.
REVIEW RETURNED	14-Aug-2018

GENERAL COMMENTS	The authors have sought to determine if there is a link between clinical time in speciality placements as an undergraduate with career destination of graduates who entered CT1/ST1 directly after FY2. This is a topical and important area to study and the authors have used innovative data to attempt to answer the question. The methods appear sound, the paper is generally well written and the conclusions contextualised. Major comments The key point to highlight in the paper is – as you acknowledge briefly in your limitations – you are only studying graduates who have entered ST1/CT1 directly from FY2 and as you point out, this is only half of graduates. This should be noted in the summary of strengths and limitations, and every time you mention the graduates entering speciality training it should state "entering speciality directly from FY2" – e.g. line 87 in abstract, line 150 in methods, line 360 in conclusion, line 367 etc. You have done this only in the title of figure 5. You data on number of clinical weeks in general practice does not correlate with other published studies (Harding, Alberti) and this should be acknowledged and discussed – it may be due to your 3rd limitation (excluding SSCs) and could this explain your lack of a statistical correlation compared to Alberti et al? You should also acknowledge that many rotations/placements are becoming integrated (e.g. LICs) and likewise these may be difficult to "label" under a speciality. The paper needs a statistician to review the methods used. They need to explain for non-statisticians whether the model could include some specialties having a strong link and others having no link at all? I don't think you can say "all" UK medical schools (strengths and limitations on page 4) when you have excluded some of them and in particular did not have data from 2 of the 30 schools you intended to include. "All" specialties may also be too strong to assert given you excluded the outliers. Similar comment to "every medical school" in line 319. Minor comments 2nd sentence of introduction is unclear Line 233 –you can say they did not allocate any time solely to histopathology and labelled as such, not that there was no time allocated Line 275 – similar point to line 275 - specific clinical time "and labelled as such" Line 301 – your argument that only the clinical placement time in GP was associated with career destination in a previous study does not really hold given that that was the definition of exposure you also used? Reference 12 – this is a letter/viewpoint not original data and a better source (e.g. GMC data) should be used Line 356 – better to state it does not appear "from our study" Line 150 – "ultimately" being appointed suggests that you have collected long term data whereas you in fact only included data of those going straight into ST/CT

VERSION 1 – AUTHOR RESPONSE

Reviewer: 1

Reviewer Name: Charles Weissman

Institution and Country: Hadassah Hebrew University Medical Center, Israel

Please state any competing interests or state 'None declared': None

● This manuscript examines the relationship of time spent in the various clinical rotations during medical school and the residency programs the students entered.

● 1. The premise of this manuscript is somewhat simple. Namely, that the amount of clinical exposure students get in medical school will direct them to select the specialty for residency, i.e. the more tthe exposure the more likely the students will select the specialty as a career. Is this true? It is much more complicated, since it is also the quality of the experience and the presence of influential mentors and the enthusiasm generated toward the students and the specialty during their rotation which are influential in specialty selection (there are many papers written on the subject of mentorsip, positive/negative experiences and specialty selection). Furthermore, personal and socio-economic factors play important roles in specialty selection and were not examined in this study.

Thank you for your comments. We agree that our manuscript starts with a simple premise and is looking at just one possible factor that may influence specialty selection. However, we feel that rather than a weakness this is in fact a strength of our study and our focus was quite deliberately on this point. We have edited our paper to further acknowledge some other possible factors (lines 244-254), and explain why we believe our study is a valuable and unique contribution to the current body of data.

● 2. The role of the medical school is to provide all graduates with a good foundation in clinical medicine no matter which specialty the student will enter. Therefore, exposure during the clinical rotations are geared to providing the student with a "bread and butter" experience. It is not the main mission of the medical school to direct students to specialties with workforce deficits. The medical school can help direct students by providing required exposures to these specialties but not at the expense of their core education. Tis issue needs to be addressed in the manuscript: Educational goals versus operational (healthcare system) goals.

Thank you for the comment. The goal of our paper was not to judge or determine the role of the medical school. However, as we have discussed, workforce requirements undoubtedly influence curriculum design. We have elaborated further on the many factors that can influence curriculum design in the Discussion of our paper (lines 484-496, lines 539-547)

● 3. If there is such a need for General Practitioners in the UK it is the obligation of the healthcare system to examine the problem and develop solutions as to why there is a problem. The medical schools can only be part of a comprehenisve solution since medical students are very savy as to the working conditions and remuneration of the various specialties. The authors did not examine the students' percentions, goals and decision making to examine the disconnect between the time spent in a specialty in medical school and their choices of specialties. This is a major weakness of this manuscript. What does the student think?

A key strength of our study, we believe, is the large data set of graduates we have looked at, however the nature of this data makes it impractical to simultaneously explore individual's goals and perceptions. We found and extensively cited previous research on student perceptions and the factors influencing their decision making, some of which has suggested that exposure during medical school is an one factor that could be easily increased to increase recruitment (see lines 246-250, lines 261-

266). It was not the aim of our study to survey students' opinions in the way suggested because numerous previous studies have already done so. Our aim was to utilise an objective, rather than subjective, data source to ascertain whether the length of time spent on a specialty appeared to influence graduates' choice of that specialty and thus provide a novel contribution to the literature on this subject.

- 4. How have other countries dealt with these issues, specialties with workfore shortages versus medical school curriculae? Research has been performed on recruiting more students to general/family practice and to rural areas by gearing medical school curriculae and experiences to shed a positive light on this area of medical practice. These studies should be reviewed in the manuscript.

Thank you for the suggestion, this has been added (line 247-250).

- 5. The authors include the surgical experience during medical school in their analysis, was it all general surgery or also exposure to surgical subspecialties? How was a student choosing urology or plastic surgery considered?

As clarified in the text of the Methods (lines 322-327) and Appendix Table A1, we grouped all surgical subspecialties into "Surgery" for this analysis. As mentioned in the Methods, this was for two reasons: firstly, several medical schools did not provide information on every subspeciality, or did not mandate students to be placed in a range of subspecialties. Secondly, the surgical subspecialties that do recruit after F2 (cardiothoracic, neurosurgery, maxillofacial surgery) also hold ST3 recruitment rounds, after two years of Core Surgical Training.

Reviewer: 2

Reviewer Name: Katherine Woolf

Institution and Country: UCL, UK

Please state any competing interests or state 'None declared': Educational advisor MRCP(UK), National Institute for Health Research Fellow.

- This paper has the potential to make a significant contribution to the literature on medical school training and specialty outcomes. I was particularly impressed at the efforts made to collect data.

The paper needs much more clarity on the methods and results. Some of the statistical analysis may need redoing since I'm not certain that the non-statistically significant findings are not due to lack of power and/or outliers. Furthermore, I could not understand one of the analyses.

- Thank you for your detailed comments and feedback, we have responded to your concerns below.

Major comments:

- 1. The interpretation of the non-significant correlation between median number of weeks spent in each specialty and number of training posts is problematic. Because the analysis is done at the level of the specialty, there are only 12 data points (as can be seen in Fig 4). The correlation coefficient is 0.43, but with an alpha of 0.05 and a sample size of 12, the power is only 0.27 (calculations done with this online calculator https://www.anzmtg.org/stats/PowerCalculator/PowerCorrelation). I personally don't think statistical significance is particularly relevant here. To me the interesting thing is that GP is such a clear outlier,

with many more posts than median weeks across medical schools. If we are going to think about a statistical relationship, the authors need to consider that such a large outlier is likely to significant skew the correlation, particularly with a small sample size. My guess from looking at the numbers in Fig 4 is that if GP were removed, the correlation may be much higher. One suggestion would be for the authors to keep Figure 4, but plot two fit lines, one including GP and one excluding it. This would allow the reader to have a visual representation of the specialties with relatively more weeks per post than average, and fewer weeks per post than average. I would not necessarily calculate the correlation; however if the authors really want to do it, I think they should remove GP as a very significant outlier and/or calculate a non-parametric correlation coefficient.

○ Thank you for this observation. We agree with your comments.

○ We have now engaged senior professor of statistics (Prof Max Bulsara) and have made important changes with his advice and as per your suggestions.

○ We have removed the p values and calculations of statistical significance, and have instead described the data.

○ We have plotted an alternative version of Figure 4 with a line of best fit that excludes GP, in order to allow the reader to quickly visualise those specialties with fewer or greater weeks of training than average.

● 2. I cannot understand how the analysis looking at the relationship between the number of weeks per specialty and the percentage of entrants per specialty for each medical school was done. I can't understand what the dependent variable is. On p12 in the results section, I don't understand the sentence on line 252-253: what does "based on a univariate analysis of variance of our dataset" mean? How was specialty "a confounder"? More information is needed on p10 in the methods.

This needs to be rewritten with more details before I can assess whether the analysis was appropriate, and how to interpret the findings.

○ We have completely re-written this section in the Results to make it clearer.

○ The Dependent variable is the percentage of F2 alumni choosing a particular specialty. The independent variables were the number of weeks spent on that specialty in medical school, the specialty itself, and the medical school itself.

○ We fitted a general linear model controlling for specialty, number of weeks spent in that Speciality, and medical school, and we report on the mean percentage of doctors entering that specific speciality after FY2.

● 3. In general the authors need to ensure consistency in the terms use to describe variables throughout the manuscript. On p10 it's stated that one of the variables is "the number of applicants from that medical school applying to each of the CT1/ST1 specialties". But then on p12 it is stated that the variable is in fact the "percentage of graduates from a medical school picking a specialty" (which is more appropriate considering medical schools produce different numbers of specialties). However on p8 it appears that these data are actually entrants to a specialty ("doctors directly entering a specialty training programme after foundation training") rather than applicants. We have clarified that we had meant to use percentages of medical school graduates who entered a specialty. Another example: on p7 an aim is to "examine whether the percentage of time spent in the different specialties correlated with the number of posts available at CT1/ST1". Then on p10 it is stated that "linear correlation was used to compare the median weeks spent in a specialty at medical school with […] the number of CT1/ST1 posts in 2016". The number of weeks spent is not the same as the percentage of time spent. We have fixed this to consistently refer to the number of weeks or absolute amount of time rather than percentages. On p12 the authors write "we found no correlation between

the number of applicants to a specialty training programme…" Do they mean applications? Or did they consider people making multiple applications to different specialties? We have changed this to "applications".

- 4. The authors need to be clearer about the level at which each analysis is done. For example, the analysis above (correlation between median weeks per specialty and number of posts) is done at the level of the specialty, i.e. not taking into account variability between medical schools. The analysis of weeks per specialty and percentage of graduates entering the specialty is presumably done at the level of the medical school?

  - We have made this clearer throughout the methods and results by specifying the level of analysis in both the titles and the main body text.

Minor comments:

- Descriptive statistics for each variable are needed. Figure 2 should be number in the text. I would find it helpful to have a table with the number and proportion of graduates per specialty per medical school.

  - Thank you - we have tried to clarify by adding descriptive statistics in the Results section (411-439).

- Those not familiar with UK medical training would probably find it helpful to have a description, and an explanation of what ST1/CT1 means.

  - We have added this explanation to the Introduction of our manuscript (lines 235-242)

- The authors asked about weeks during "clinical education", however in some schools the division between pre-clinical and clinical is not clear, in particular some less traditional schools have quite a lot of early patient contact. This needs to be mentioned.

  - Thank you for the suggestion, we have added this to the discussion (line 586-593).

Reviewer: 3

Reviewer Name: Hugh Alberti

Institution and Country: Newcastle University, UK

Please state any competing interests or state 'None declared': None declared. I was the author of one of the papers that the authors quote and discuss.

- The authors have sought to determine if there is a link between clinical time in speciality placements as an undergraduate with career destination of graduates who entered CT1/ST1 directly after FY2. This is a topical and important area to study and the authors have used innovative data to attempt to answer the question. The methods appear sound, the paper is generally well written and the conclusions contextualised.

  - Thank you.

Major comments

- The key point to highlight in the paper is – as you acknowledge briefly in your limitations – you are only studying graduates who have entered ST1/CT1 directly from FY2 and as you point out, this is only half of graduates. This should be noted in the summary of strengths and limitations, and every

time you mention the graduates entering speciality training it should state “entering speciality directly from FY2” – e.g. line 87 in abstract, line 150 in methods, line 360 in conclusion, line 367 etc. You have done this only in the title of figure 5.

○ Thank you for the suggestion, we have reviewed the entire manuscript and now clarified this in the abstract, methods, results, discussion, and conclusion. We have also better highlighted the source of our data (the mandatory National F2 Career Destination survey), and therefore explained the reason why we could only access data on this half of graduates (see line 303-306). We have also amended the limitations section (lines 569-570).

● You data on number of clinical weeks in general practice does not correlate with other published studies (Harding, Alberti) and this should be acknowledged and discussed – it may be due to your 3rd limitation (excluding SSCs) and could this explain your lack of a statistical correlation compared to Alberti et al? You should also acknowledge that many rotations/placements are becoming integrated (e.g. LICs) and likewise these may be difficult to “label” under a speciality.

The paper needs a statistician to review the methods used. They need to explain for non-statisticians whether the model could include some specialties having a strong link and others having no link at all?

○ We have added a paragraph in the discussion to acknowledge this point and discuss general practice further.

○ We have also involved a Professor of Statistics to review the methods and help explain the methods and results in a clearer way.

● I don’t think you can say “all” UK medical schools (strengths and limitations on page 4) when you have excluded some of them and in particular did not have data from 2 of the 30 schools you intended to include. “All” specialties may also be too strong to assert given you excluded the outliers. Similar comment to “every medical school” in line 319.

○ We have changed the text to make this clearer throughout.

Minor comments

● 2nd sentence of introduction is unclear

○ This has been amended (line 229).

● Line 233 –you can say they did not allocate any time solely to histopathology and labelled as such, not that there was no time allocated

○ This has been amended (line 389).

● Line 275 – similar point to line 275 - specific clinical time “and labelled as such”

○ This has been amended (lines 484).

● Line 301 – your argument that only the clinical placement time in GP was associated with career destination in a previous study does not really hold given that that was the definition of exposure you also used?

○ Thank you - we feel it is possible that our method of data collection may also have included specialty exposure during clinical years that did not involve direct patient contact. Additionally, we were looking at a different cohort (2 years later). Finally, the association found in the previous study was relatively weak. We have elaborated on this in lines 515-537 in the Discussion.

● Reference 12 – this is a letter/viewpoint not original data and a better source (e.g. GMC data) should be used

○ Thank you for your comment. We have edited the sentence this reference refers to and provided a new reference accordingly (lines 265-268).

● Line 356 – better to state it does not appear “from our study”

○ The conclusion has been amended and we have clarified that analyses from our study do not suggest this (lines 607-609).

● Line 150 – “ultimately” being appointed suggests that you have collected long term data whereas you in fact only included data of those going straight into ST/CT

○ Thank you for pointing this out, we have removed it and clarified we mean only those directly entering CT1/ST1 (line 284-286).

VERSION 2 – REVIEW

REVIEWER	Charles Weissman Hadassah-Hebrew University Medical Center, Israel
REVIEW RETURNED	05-Dec-2018

GENERAL COMMENTS	Much improved

REVIEWER	Hugh Alberti School of Medical Education Newcastle University Newcastle UK
REVIEW RETURNED	I was the author of one of the papers quoted in the paper

GENERAL COMMENTS	The authors have responded to the majority of my previous points and the paper is now stronger. However not all points were addressed and there are some issues with new sections of the paper: Line 93 – what do you mean by jobs here? Available training posts? Line 106 = appointed directly to Line 121 – you looked at career training allocations not decisions Line 134 – need to add why this is an issue – i.e. whereas the curriculum time data was collected in… and many of the curricula may have changed since the doctors in the study were students Line 203 –need to add authentic or clinical (or both) as there was not an association with overall time, just with the clinical/authentic time Line 204 – need to add directly Line 377 – need to add directly from F2 Line 494 - directly I also note the considerable concerns raised by reviewer 2 regarding the statistics in the paper and would recommend that it is only published if all of these concerns are addressed to the satisfaction of reviewer 2 or a statistician. I would also recommend the editor reviews the contribution made by the additional author to the revision to consider if their input equates with authorship.

REVIEWER	Dr. Zoe Hoare NWORTH CTU, Bangor University UK
REVIEW RETURNED	12-Feb-2019

GENERAL COMMENTS	The study describes a retrospective observational study of the relationship between time spent training in a speciality during undergraduate training and the applications to the associated specialities later. The analysis and research question appears to define the question/hypothesis as a mechanism that works in a forwards direction that exposure to results in application to. However is it not more likely that demand in the system generates the training need? The authors rightly identify the limitation of the work in relation to the time lag between the applications and the data collection of course descriptions. It may be useful to include at least some commentary on the variability of courses year on year and the variability of the specialty availability year on year. I would imagine that demand for various reasons varies widely in terms of speciality - although it is likely there will always be a greater number of GPs needed than more highly focussed specialities; surely this is a system driven demand that can only be retrospectively reacted to in terms of undergraduate course development. Is it possible to take into account and develop whether there is any impact in the opposite direction to that hypothesised? I doubt there is with the current dataset collected. The general linear model appears to be underspecified (or at least limited in description) were there no other contributory factors that were considered? No sensitivity analyses were performed in relation to the full model (aside from excluding GPs - were they included in the modelling or not?) it would be useful to know the impact of the inclusion of the unknown category on the model as Figure 3 indicates that this could be a contributor of noise to the model and is poorly defined as it could incorporate many of the other specialty aspects considered as separate categories. Whilst the relevant coefficient for the model is quoted in the text it would be helpful to include the model definition for clarity. The time spent in speciality training and the number of weeks in clinical training are described but the relationship between these elements is not described - do those with the most time in clinical training also have the most time in speciality training? Does the clinical training factor impact on the model regression model defined in anyway? Description of the models dependent variable switches between number entering specialty training and percentage this needs to be consistent throughout in terms of which one is being used. I think some further explanation of the analysis methodology and additional sensitivity analysis may stregthen the work but I am concerned that the underlying premise is why we would expect a relationship in the direction noted?

VERSION 2 – AUTHOR RESPONSE

Reviewer: 1

Reviewer Name: Charles Weissman

Institution and Country: Hadassah-Hebrew University Medical Center, Israel

Please state any competing interests or state 'None declared': none

Please leave your comments for the authors below

- Much improved
    - Thank you for your comments.

Reviewer: 3

Reviewer Name: Hugh Alberti

Institution and Country: School of Medical Education - Newcastle University - Newcastle - UK

Please state any competing interests or state 'None declared': I was the author of one of the papers quoted in the paper

Please leave your comments for the authors below

- The authors have responded to the majority of my previous points and the paper is now stronger.
    - Thank you for your comments and recommendations.
- However not all points were addressed and there are some issues with new sections of the paper
- Line 93 – what do you mean by jobs here? Available training posts?
    - Yes - number of available training posts. Edited to clarify.
- Line 106 = appointed directly to
    - Added
- Line 121 – you looked at career training allocations not decisions
    - Amended
- Line 134 – need to add why this is an issue – i.e. whereas the curriculum time data was collected in… and many of the curricula may have changed since the doctors in the study were students
    - Added and clarified.
- Line 203 –need to add authentic or clinical (or both) as there was not an association with overall time, just with the clinical/authentic time

    - Sentence edited to add “clinical” before “GP training” (line 154).

- Line 204 – need to add directly
    - Added (line 155)
- Line 377 – need to add directly from F2
    - Added (line 349)
- Line 494 - directly
    - Added (line 421)
- I also note the considerable concerns raised by reviewer 2 regarding the statistics in the paper and would recommend that it is only published if all of these concerns are addressed to the satisfaction of reviewer 2 or a statistician. I would also recommend the editor reviews the contribution made by the additional author to the revision to consider if their input equates with authorship.
    - The additional author (Max Bulsara) is a professor of biostatistics at the University of Notre Dame, adjunct professor at the UWA School of Public Health, and a visiting professor at University College London. We had engaged with Prof. Bulsara before the initial submission informally, and he had helped guide us with making the general linear model that comprises a large part of the methods and results of our paper. We reached out to him again after the first review process in order to help clarify and amend our methods, and ensure a rigorous statistical approach throughout the entire paper. His input has proved invaluable in addressing the concerns and comments raised by the reviewers, as well as improving our methods and clarifying our results. All authors agree Prof Bulsara fulfils the four criteria ICMJE recommend for authorship.
    - We have been informed that reviewer 2 had no further comments during this round of reviews.

Reviewer: 4

Reviewer Name: Dr. Zoe Hoare

Institution and Country: NWORTH CTU, Bangor University - UK

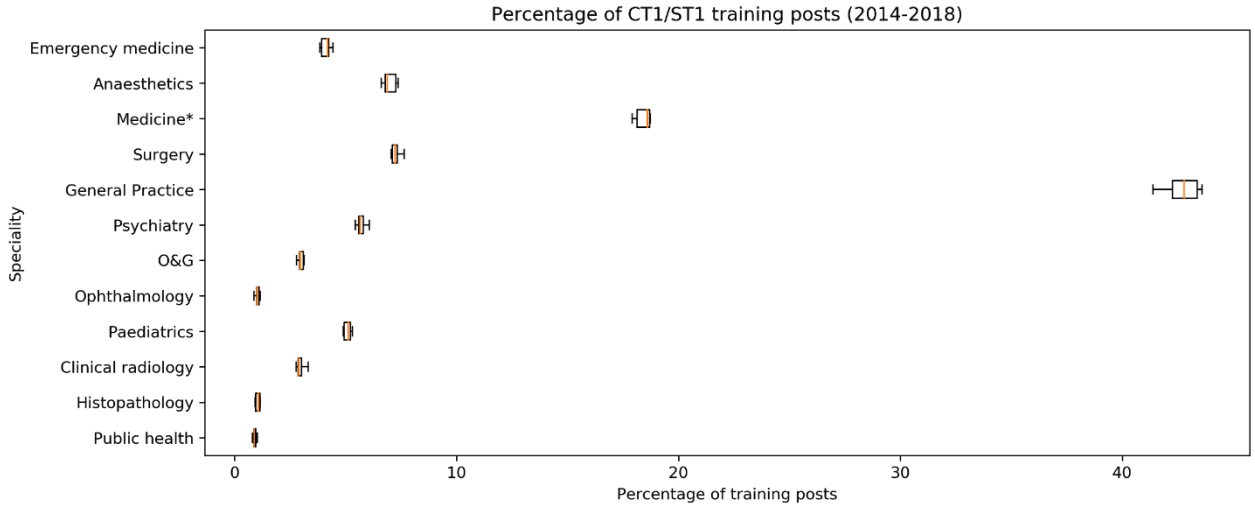
Please state any competing interests or state ‘None declared’: None declared

Please leave your comments for the authors below

- The study describes a retrospective observational study of the relationship between time spent training in a speciality during undergraduate training and the applications to the associated specialities later.
- The analysis and research question appears to define the question/hypothesis as a mechanism that works in a forwards direction that exposure to results in application to. However is it not more likely that demand in the system generates the training need?
    - Thank you for your comments. We think that clinical service requirements and patient demand will likely generate training need. Currently, there are recruitment issues as

some specialties (such as GP or A&E) are struggling to find doctors. One of the strategies to tackle this problem may be to increase exposure to those specialties during medical school. However, this is based on the premise that increased exposure would increase the proportion of doctors that would choose a specialty. This is indeed supported by previous literature on GP exposure. If this is true, one would expect that those medical schools that currently have greater exposure to any given specialty would have a greater proportion of their graduates entering that specialty. We tested this by defining a general linear model that incorporated two independent variables (X1, X2) and one dependent variable (Y): (X1) the number of weeks spent in a specialty at medical school (X2) the specialty, and (Y) the percentage of graduates from that medical school that entered each specialty.

- The authors rightly identify the limitation of the work in relation to the time lag between the applications and the data collection of course descriptions. It may be useful to include at least some commentary on the variability of courses year on year and the variability of the specialty availability year on year. I would imagine that demand for various reasons varies widely in terms of speciality - although it is likely there will always be a greater number of GPs needed than more highly focussed specialities; surely this is a system driven demand that can only be retrospectively reacted to in terms of undergraduate course development. Is it possible to take into account and develop whether there is any impact in the opposite direction to that hypothesised? I doubt there is with the current dataset collected.
    - We have now clarified this limitation further in our Limitations section.
    - We do not have data on the variability of medical school courses year on year, and so we do not think it would be possible to see whether variability in job availability influences medical school curricula.
    - It is possible to examine the variability of specialty training posts year on year. Based on the published data on specialty training posts and applications from Health



Education England, we have produced a box-plot, which we will include as a Appendix figure (2). The year-on-year variability is small.

- The general linear model appears to be underspecified (or at least limited in description) were there no other contributory factors that were considered?
    - The only factors we considered in our general linear model were the specialty in question and the number of weeks spent on the specialty. This may be visualised by figure 5, where our two factors were the number of weeks (X-axis) and the specialty

(colour/shape of the marker). We only fitted the GLM with variables that were available to us, so we understand the model is underspecified. However, we are not using this model for predictive purposes - it is used to explore significant factors related to our outcome measure.

- No sensitivity analyses were performed in relation to the full model (aside from excluding GPs - were they included in the modelling or not?) it would be useful to know the impact of the inclusion of the unknown category on the model as Figure 3 indicates that this could be a contributor of noise to the model and is poorly defined as it could incorporate many of the other specialty aspects considered as separate categories.
    - We did not exclude GP from the general linear model; it was only excluded for our initial regression. We have now clarified this in the text, and have added the relevant statistics for this regression.
    - In our general linear model, we excluded Unknown category as including it would just add noise to the model, and reduce our sensitivity.
- Whilst the relevant coefficient for the model is quoted in the text it would be helpful to include the model definition for clarity.
    - We include below a table of the beta coefficients to help define the model. If desired we can include this as a figure in the appendix.

	Coefficient	P
Weeks	0.061	0.228
ACCS	0 (base)	
General Practice	10.4	0.000
Medicine	3.25	0.011
Obstetrics & Gynaecology	-4.42	0.000
Ophthalmology	-0.48	0.000
Paediatrics	-2.82	0.000
Psychiatry	-3.32	0.000
Surgery	-1.13	0.184
Constant	5.67	0.000
Aberdeen	0 (base)	
Barts - London	-0.22609	0.866
Birmingham	0.032191	0.981
Brighton & Sussex	-0.30539	0.821
Bristol	-1.60121	0.232
Cambridge	1.217413	0.364
Cardiff	-1.23189	0.358
Dundee	-0.82189	0.54
Edinburgh	-1.32505	0.324
Glasgow	-1.25159	0.35
Hull York	0.188431	0.889
Imperial College London	-0.43789	0.745
Keele	0.932636	0.486
King's College London	-0.3372	0.802

Lancaster	0.78023	0.56
Leeds	-1.98348	0.139
Leicester	1.143159	0.397
Liverpool	0.118273	0.93
Norwich	0.206598	0.877
Nottingham	0.469954	0.727
Oxford	0.640609	0.634
Plymouth	-0.61673	0.645
Queen's Belfast	-1.08798	0.417
Sheffield	-0.02279	0.986
Southampton	-0.68501	0.609
St George's London	-0.15873	0.906
UCL	-1.31541	0.326

- The time spent in speciality training and the number of weeks in clinical training are described but the relationship between these elements is not described - do those with the most time in clinical training also have the most time in speciality training? Does the clinical training factor impact on the model regression model defined in anyway
    - When we refer to ‘medical specialties’ we are referring to all clinical pathways or specialities, (including eg. GP). We measured the time spent in “specialty training” during medical school in weeks, and this term is equivalent to the time in “clinical training”.
- Description of the models dependent variable switches between number entering specialty training and percentage this needs to be consistent throughout in terms of which one is being used.
    - We have tried to make this correct and consistent throughout. In the first part of our analysis (Figure 4) we use the number of specialty training posts, as we did this analysis considering nationwide training posts. However, for our model (& Figure 5) we were considering individual medical schools, and so we used a percentage value, based on the percentage of each medical school’s graduates who entered a specialty. Since medical schools differ widely in size, using the total numbers at this point would have been inappropriate.
- I think some further explanation of the analysis methodology and additional sensitivity analysis may stregthen the work but I am concerned that the underlying premise is why we would expect a relationship in the direction noted?
    - Thank you for your considered feedback on our paper, we hope our comments above help to clarify your concerns.

VERSION 3 – REVIEW

REVIEWER	Dr. Zoe Hoare NWORTH, Bangor University UK
REVIEW RETURNED	16-May-2019

GENERAL COMMENTS	The authors have addressed all of the reviewers comments well. The manuscript now reads well and the analysis section is much clearer.

REVIEWER	Hugh Alberti Newcastle University United Kingdom Only that I am the author of one of the papers that is discussed in this paper.
REVIEW RETURNED	22-May-2019

GENERAL COMMENTS	I am happy with the further corrections that the authors have made to the paper